



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2011

Schlussprüfungen an der Weiterbildungsschule Basel-Stadt. Schlussbericht 2011 zuhanden der Stufenleitung Sekundarstufe I des Kantons Basel-Stadt

Keller, F

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-59972>

Published Research Report

Published Version

Originally published at:

Keller, F (2011). Schlussprüfungen an der Weiterbildungsschule Basel-Stadt. Schlussbericht 2011 zuhanden der Stufenleitung Sekundarstufe I des Kantons Basel-Stadt. Zürich: Institut für Bildungsevaluation, Universität Zürich.



**Universität
Zürich**^{UZH}

**Institut für Bildungsevaluation
Assoziiertes Institut der Universität Zürich**

Schlussprüfungen an der Weiterbildungs- schule Basel-Stadt

**Schlussbericht 2011 zuhanden der Stufenleitung Sekundarstufe I des
Kantons Basel-Stadt**

Florian Keller

Zürich, Juli 2011



**Universität
Zürich**^{UZH}

Institut für Bildungsevaluation
Assoziiertes Institut der Universität Zürich
Wilfriedstrasse 15
8032 Zürich

Tel: 043 268 39 60
Fax: 043 268 39 67

www.ibe.uzh.ch

Inhaltsverzeichnis

1 Einleitung	4
2 Das Wichtigste in Kürze	5
3 Fachleistungen der Schülerinnen und Schüler	8
3.1 Mathematikleistungen am Ende der 9. Klasse	8
3.2 Deutschleistungen am Ende der 9. Klasse	11
3.3 Leistungsunterschiede zwischen Mädchen und Knaben am Ende der 9. Klasse	14
4 Fachleistungen nach Klassen	15
4.1 Mathematik- und Deutschleistungen nach Klassen	15
4.2 Beurteilung der Unterschiede zwischen den Klassen	17
5 Fachleistungen nach WBS-Standorten	18
6 Leistungsentwicklung an der WBS	19
6.1 Entwicklung der durchschnittlichen Leistungen der Schülerinnen und Schüler an der WBS	19
6.2 Entwicklung der Leistungsunterschiede zwischen Mädchen und Knaben	21
6.3 Entwicklung der Fähigkeiten der Schülerinnen und Schüler an der WBS	22
7 Leistungsentwicklung an der WBS seit der Umsetzung des «Mathe-Pakets»	24
8 Fazit	31
Anhang	32
Glossar der statistischen Begriffe	32
Daten zu den Abbildungen 6.1 bis 6.3	33

1 Einleitung

Am Ende der obligatorischen Schulzeit werden an der Weiterbildungsschule Basel-Stadt (WBS) Schlussprüfungen durchgeführt. Mit den Schlussprüfungen werden die Fähigkeiten der Schülerinnen und Schüler in den Fächern Deutsch und Mathematik mit einem einheitlichen Leistungstest geprüft und benotet. Die Noten sind Teil des Schlusszeugnisses der Schülerinnen und Schüler.

Die Schlussprüfungen werden seit 2007 vom Institut für Bildungsevaluation (IBE) durchgeführt. Das IBE übernimmt im Auftrag der Stufenleitung Sekundarstufe I der Volksschule des Kantons Basel-Stadt die Entwicklung und Erprobung der Prüfungsaufgaben sowie die Korrektur, Auswertung und Rückmeldung der Prüfungsergebnisse. Nach Abschluss der Schlussprüfungen werden die Prüfungsergebnisse in einem Bericht dargestellt und mit den bisherigen Ergebnissen verglichen. Dadurch dienen die jährlichen Berichte zu den Schlussprüfungen der Stufenleitung Sekundarstufe I auch als ein Instrument der Qualitätssicherung.

Im Zentrum der bisherigen Schlussberichte standen die Mathematikleistungen an der WBS. Bereits im ersten Schlussbericht 2007 wurde auf die schwachen und stagnierenden Mathematikleistungen an der WBS hingewiesen. Aufgrund dieser Resultate wurden auf das Schuljahr 2008/09 unter dem Begriff «Mathe-Paket» verschiedene Massnahmen zur Steigerung der Mathematikleistungen ergriffen. Unter anderem wurde zu Beginn der 1. Klasse der WBS eine Lernstandserhebung durchgeführt und die Zahl der wöchentlichen Mathematikektionen wurde um eine Lektion erhöht. Dieses Jahr wurden die Schlussprüfungen zum zweiten Mal bei Schülerinnen und Schüler durchgeführt, die während ihrer gesamten Schulzeit an der WBS von den leistungssteigernden Massnahmen profitieren konnten. Damit wird es möglich, erste statistisch fundierte Antworten auf die Frage zu geben, ob das «Mathe-Paket» zu einer Steigerung der Mathematikleistungen an der WBS geführt hat.

Der Bericht ist in acht Kapitel gegliedert. In Kapitel 2 werden die Durchführung der Schlussprüfungen und die Berechnung der Leistungsskalen beschrieben. Zudem werden die wichtigsten Ergebnisse der Schlussberichte der letzten Jahre zusammengefasst. In Kapitel 3 werden die Leistungen sowie die Fähigkeiten der Schülerinnen und Schüler in den Schlussprüfungen 2010 beschrieben. In Kapitel 4 sind die durchschnittlichen Leistungen der Klassen und in Kapitel 5 die durchschnittlichen Leistungen der einzelnen WBS-Standorte in den Schlussprüfungen 2010 dargestellt. In Kapitel 6 werden die diesjährigen Prüfungsergebnisse mit den Ergebnissen der Jahre 2005 bis 2010 verglichen und die Leistungsentwicklung an der WBS aufgezeigt. Die Frage, ob das «Mathe-Paket» eine leistungsrelevante Wirkung hat, wird in Kapitel 7 beantwortet. Der Bericht schliesst mit einer Zusammenfassung der wichtigsten Resultate.

2 Das Wichtigste in Kürze

Testinstrumente – Die Testinstrumente, die als Schlussprüfungen am Ende der 2. Klasse der WBS eingesetzt werden, basieren auf Leistungstests, die 2004 im Rahmen der Evaluation der Strukturänderung an der WBS am Institut für Bildungsevaluation entwickelt wurden. Seither werden die Testaufgaben jedes Jahr überarbeitet und verbessert. Damit die Testresultate gleichwohl vergleichbar bleiben, werden ausgewählte Aufgaben als «Link-Items» während mehrerer Jahre identisch eingesetzt.

Der Mathematiktest 2011 umfasst 75 Aufgaben aus den Lehrplanbereichen «Zahl und Zahloperationen», «Funktionen», «Gleichungen und Ungleichungen», «Geometrie», «Sachrechnen» und «Elementare Statistik».

Der Deutschttest 2011 umfasst 112 Aufgaben aus den Lehrplanbereichen «Lesen und Verstehen», «Grammatik» und «Rechtschreibung». Der Lehrplanbereich «Schreiben» wurde mit einem Schreibanlass geprüft. Drei Themen standen dabei zur Auswahl: «Mein schönstes Schulerlebnis», «Was ich von einer guten Freundin/einem guten Freund erwarte» und «Das Handy – Fluch oder Segen unserer Zeit?».

Die Tests sind so konzipiert, dass alle Schülerinnen und Schüler unabhängig vom besuchten Klassenzug die gleiche Prüfung lösen und dass die Prüfungsergebnisse mit dem Referenzrahmen der Evaluation der Strukturänderung an der WBS (Moser & Keller, 2006) empirisch verglichen werden können.

Teilnahme – An den Schlussprüfungen 2011 nahmen 802 Schülerinnen und Schüler aus 47 Klassen teil. 313 Schülerinnen und Schüler (39 Prozent) wurden in einer Regelklasse des A-Zugs, 410 Schülerinnen und Schüler (51 Prozent) in einer Regelklasse des E-Zugs unterrichtet. 81 Schülerinnen und Schüler (rund 10 Prozent) besuchten eine Musikklasse, eine Sportklasse oder die Fremdsprachenklasse.

Durchführung – Die Schlussprüfungen fanden im Juni 2011 statt und wurden von den sechs Weiterbildungsschulen selbstständig organisiert. Die Prüfungshefte wurden anschliessend ans Institut für Bildungsevaluation nach Zürich gebracht, wo sie von einem Team von Fachwissenschaftlern mit Unterrichtserfahrung korrigiert wurden. Die Aufsätze des Deutschttests (Lehrplanbereich «Schreiben») wurden mittels eines standardisierten Korrekturrasters nach zehn inhaltlichen und zehn formalen Kriterien bewertet. Nach der Korrektur wurden die Aufgaben elektronisch erfasst und plausibilisiert. Die Testresultate der Schülerinnen und Schüler wurden anhand eines Notenschlüssels, der seit 2007 in nahezu unveränderter Form eingesetzt wird, in Noten umgerechnet.

Zwei Wochen nach der Durchführung der Schlussprüfungen konnten die Noten der Schülerinnen und Schüler an die Schulleitungen versandt werden. Den Lehrpersonen wurden die Prüfungsergebnisse ihrer Schülerinnen und Schüler als prozentualer Anteil richtig gelöster Aufgaben sowie das durchschnittliche Klassenergebnis zugestellt. Zudem erhielten die Schulleitungen einen kurzen Bericht zu den durchschnittlichen Leistungen der Klassen an ihrer Schule und für die Stufenleitung wurde eine Rückmeldung mit den Resultaten der WBS-Standorte erstellt.

Skalierung der Leistungsdaten – Die Leistungen der Schülerinnen und Schüler in den Schlussprüfungen werden als Punktzahl auf der WBS-Skala ausgewiesen. Die Punktzahl wurde mit einer Rasch-Skalierung unter Einbezug der Parameter der Schlussprüfungen 2005 berechnet. Diese Methode ermöglicht es, die Leistungen der Schülerinnen und Schüler in den Jahren 2005 bis 2011 trotz teilweise neuer Aufgaben als Punktzahl auf der gleichen Skala zu verorten. Die Skala ist so standardisiert, dass der Mittelwert aller Schülerinnen und Schüler der Schlussprüfungen 2005 in jedem Fach 500 Punkte und die Standardabweichung 100 Punkte beträgt. Veränderungen der durchschnittlichen Schülerleistungen sowie die Leistungsentwicklung an den einzelnen Schulen können somit einfach interpretiert werden: Werte über 500 Punkte entsprechen einer höheren Leistung im Vergleich zu den Schlussprüfungen 2005, Werte unter 500 Punkte einer tieferen Leistung. Leistungsunterschiede – beispielsweise zwischen Mädchen und Knaben oder zwischen zwei Prüfungsjahren – sind dann relevant, wenn sie 20 Punkte oder mehr betragen. Unterschiede von 80 und mehr Punkten sind sehr gross.

Ein weiterer wesentlicher Vorteil dieser Skalierung ist, dass die Leistungen der Schülerinnen und Schüler und die Schwierigkeit der Prüfungsaufgaben auf der gleichen Skala mit einer einheitlichen Metrik abgebildet werden können. Die Schwierigkeit der Aufgaben und die Leistung der Schülerinnen und Schüler stehen dabei in einer definierten Beziehung zueinander. Eine Schülerin beispielsweise, die eine Leistung von 700 Punkten erreicht, kann eine Aufgabe mit der Schwierigkeit von 700 Punkten mit einer Wahrscheinlichkeit von 62 Prozent richtig lösen. Im Wissen um diese Lösungswahrscheinlichkeit können die Schülerleistungen anhand von Testaufgaben inhaltlich beschrieben werden. Damit die Beschreibung der Leistungen übersichtlich wird, wurden ähnlich schwierige Aufgaben zu Intervallen zusammengefasst.

Intervalle – Für die inhaltliche Beschreibung wurden die Leistungen der Schülerinnen und Schüler in Intervalle von jeweils 100 Punkten auf der WBS-Skala eingeteilt. Schülerinnen und Schüler, die ein bestimmtes Intervall erreichen, können mindestens die Hälfte der Aufgaben in diesem Intervall richtig lösen. Erreicht eine Schülerin beispielsweise im Mathematiktest 450 Punkte, dann beträgt die durchschnittliche Lösungswahrscheinlichkeit für die Aufgaben des Intervalls «401 bis 500 Punkte» 62 Prozent. Das heisst, die Schülerin löst mit einer Wahrscheinlichkeit von 62 Prozent alle Aufgaben dieses Intervalls richtig lösen. Erreicht eine Schülerin 401 Punkte, dann beträgt die durchschnittliche Lösungswahrscheinlichkeit für die Aufgaben des Intervalls «401 bis 500 Punkte» noch 50 Prozent. Diese Schülerin ist in der Lage 50 Prozent der Aufgaben des Intervalls richtig zu lösen.

Schlussbericht 2007 – Im Jahr 2007 wurde zum ersten Mal ein Schlussbericht zu den Schlussprüfungen der WBS vorgelegt. Darin konnte unter anderem aufgezeigt werden, dass sich die durchschnittlichen Leistungen seit 2005 in Deutsch kontinuierlich verbessert haben. In der Mathematik hingegen stagnierten die Leistungen und viele Schülerinnen und Schüler erreichten die Ziele des Lehrplans nicht. Aufgrund dieser Resultate ergriff das Rektorat der WBS unter der Bezeichnung «Mathe-Paket» verschiedene Massnahmen zur Förderung der Mathematikleistungen.

Schlussbericht 2008 – Bereits im Schuljahr 2007/08 konnten einige kleinere Massnahmen des «Mathe-Pakets» umgesetzt werden. Wie der Schlussbericht 2008 zeigt, verbesserten sich die Ergebnisse der Schülerinnen und Schüler sowohl in Deutsch als auch in der Mathematik deutlich. Dabei haben sich nicht nur die durchschnittlichen Mathematikleistungen verbessert, auch der Anteil Schülerinnen und Schüler mit ungenügenden Grundkompetenzen konnte in der Mathematik gegenüber 2007 reduziert werden.

Schlussbericht 2009 – Der Schlussbericht 2009 machte deutlich, dass die getroffenen Massnahmen nicht zu einer messbaren Steigerung der Mathematikleistungen geführt haben. Im Gegenteil: Die Leistungen der Schülerinnen und Schüler brachen 2009 im Vergleich zum Vorjahr geradezu ein. Allerdings konnten die schwachen Leistungen teilweise durch die spezifischen Merkmale des Schülerjahrgangs 2009 erklärt werden. So war der Anteil an Schülerinnen und Schülern im A-Zug sowie die Gymnasialquote klar grösser als in den Jahren zuvor. Es traten also mehr gute E-Schülerinnen und Schüler ins Gymnasium ein, was tendenziell zu schwächeren Durchschnittsleistungen der WBS-Schülerinnen und Schüler führte. Zudem war der Anteil Mädchen im Jahr 2009 sehr gross. Da Mädchen in der Mathematik durchschnittlich schwächere Leistungen erbringen, sanken damit auch die durchschnittlichen Mathematikleistungen insgesamt.

Schlussbericht 2010 – Mit den Analysen zu den Schlussprüfungen 2010 konnte aufgezeigt werden, dass sich die Leistungen an der WBS seit 2005, trotz kleineren Leistungsschwankungen zwischen den einzelnen Testjahren, stetig verbessert haben. Allerdings sind die Leistungssteigerungen gering. Dies ist an und für sich kein ungewöhnliches Ergebnis angesichts der relativ kurzen Zeitspanne, der grossen Schülerzahl und den vielfältigen Ursachen, die die schulischen Leistungen beeinflussen. Allerdings hätte erwartet werden können, dass sich die Massnahmen des «Mathe-Pakets» deutlicher in den Leistungen der Schülerinnen und Schüler niederschlagen würden. Dies konnte jedoch bislang nicht nachgewiesen werden.

Weiterführende Informationen – Informationen zum methodischen Vorgehen bei der Skalierung der Prüfungsergebnisse sowie eine ausführliche Beschreibung der Intervalle anhand von Beispielaufgaben finden sich im Schlussbericht zur Evaluation der Strukturänderung an der WBS Basel-Stadt:

- Moser, U. & Keller, F. (2006): *Evaluation der Strukturänderung an der Weiterbildungsschule Basel-Stadt. Schlussbericht zuhanden des Erziehungsdepartements des Kantons Basel-Stadt, Ressort Schulen*. Zürich: Kompetenzzentrum für Bildungsevaluation.
<http://www.ibe.uzh.ch/publikationen/Evaluation-WBS.pdf>

Die Berichte mit den Ergebnissen der Schlussprüfungen 2007 bis 2010 können auf der Website des IBE herunter geladen werden:

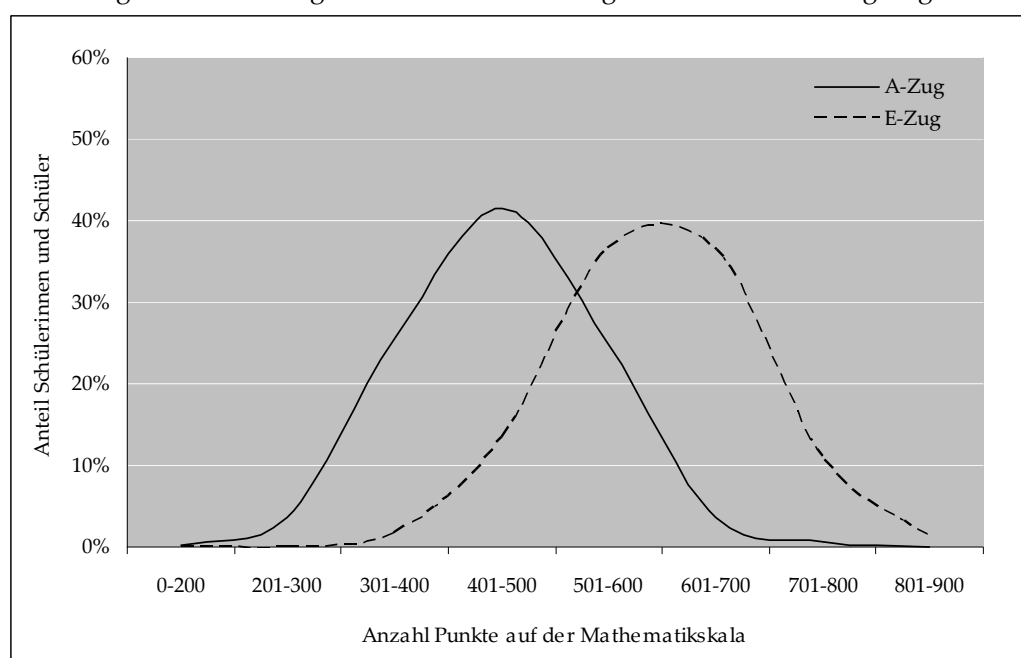
- <http://www.ibe.uzh.ch/projekte/wbsabschluss.html>

3 Fachleistungen der Schülerinnen und Schüler

3.1 Mathematikleistungen am Ende der 9. Klasse

Abbildung 3.1 zeigt die Verteilung der Mathematikleistungen der Schülerinnen und Schüler in der Schlussprüfung 2011 nach Leistungszug. Die Schülerinnen und Schüler des A-Zugs erreichen am Ende der 9. Klasse durchschnittlich 450 Punkte auf der Mathematikskala, die Schülerinnen und Schüler des E-Zugs 595 Punkte. Die Differenz zwischen den durchschnittlichen Mathematikleistungen der Schülerinnen und Schüler des A- und des E-Zugs beträgt rund 144 Punkte. Dieser Unterschied ist sehr gross (Effektgrösse $d = 1.60$).

Abbildung 3.1 Verteilung der Mathematikleistungen 2011 nach Leistungszug

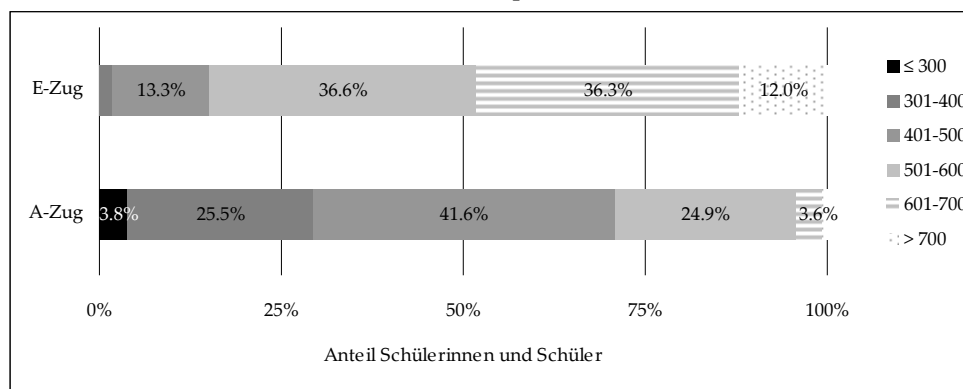


Anmerkungen: ohne Schülerinnen und Schüler der Fremdsprachenklassen
 A-Zug: N = 320; M = 450 Punkte; SD = 90 Punkte
 E-Zug: N = 457; M = 595 Punkte; SD = 90 Punkte

Trotz der grossen Unterschiede zwischen den durchschnittlichen Leistungen des A-Zugs und den durchschnittlichen Leistungen des E-Zugs überschneiden sich die Leistungsverteilungen der beiden Züge (Abbildung 3.1). Rund 42 Prozent der Schülerinnen und Schüler können aufgrund ihrer Mathematikleistungen nicht eindeutig dem A- oder dem E-Zug zugewiesen werden. Rund 4 Prozent der Schülerinnen und Schüler des A-Zugs erzielen in der Mathematik eine höhere Leistung als der Durchschnitt der Schülerinnen und Schüler des E-Zugs. Umgekehrt erreichen rund 5 Prozent der Schülerinnen und Schüler des E-Zugs tiefere Mathematikleistungen als der Durchschnitt der Schülerinnen und Schüler des A-Zugs.

Abbildung 3.2 zeigt die Verteilung der Schülerinnen und Schüler am Ende der 9. Klasse auf die Intervalle in der Mathematik nach Leistungszug. Schülerinnen und Schüler, die einem bestimmten Intervall zugeordnet werden, können die Aufgaben in diesem Intervall sowie alle Aufgaben der tieferen Intervalle mit einer Wahrscheinlichkeit von mindestens 50 Prozent richtig lösen.

Abbildung 3.2 Verteilung der Schülerinnen und Schüler auf die Intervalle in der Mathematik (ohne Fremdsprachenklassen)



29 Prozent der Schülerinnen und Schüler des A-Zugs sowie 2 Prozent der Schülerinnen und Schüler des E-Zugs erreichen am Ende der WBS weniger als 400 Punkte auf der Mathematikskala. Ihnen fehlen in der Mathematik jene Grundkompetenzen, die zur Lösung von einfachen Grundoperationen notwendig sind.

98 Prozent der Schülerinnen und Schüler des E-Zugs sowie rund 71 Prozent der Schülerinnen und Schüler des A-Zugs erreichen am Ende der WBS mehr als 400 Punkte. Sie kennen unter anderem Zahlenarten und Zahlenmengen sowie die wichtigsten Grundoperationen und können sie auch anwenden.


Rund 85 Prozent der Schülerinnen und Schüler des E-Zugs sowie rund 29 Prozent der Schülerinnen und Schüler des A-Zugs erreichen am Ende der WBS mehr als 500 Punkte. Sie können beispielsweise einfache Prozentrechnungen lösen sowie Grundoperationen mit Dezimalzahlen durchführen. Zudem können sie Umfang und Flächen von Drei- und Vierecken berechnen.

Rund 48 Prozent der Schülerinnen und Schüler des E-Zugs sowie rund 4 Prozent der Schülerinnen und Schüler des A-Zugs erreichen am Ende der WBS mehr als 600 Punkte. Sie können unter anderem Bruchgleichungen lösen und Folgerungen aus grafischen Darstellungen von statistischen Daten ziehen.

Rund 12 Prozent der Schülerinnen und Schüler des E-Zugs sowie 2 (1%) Schülerinnen und Schüler des A-Zugs erreichen am Ende der WBS mehr als 700 Punkte. Sie können unter anderem Potenzen und Quadratwurzeln anwenden sowie Rauminhalte und Flächen von geometrischen Körpern berechnen.

In Tabelle 3.1 sind für jedes Intervall in der Mathematik Beispiele typischer Aufgaben aus der WBS Schlussprüfung 2011 dargestellt.

Tabelle 3.1 Aufgabenbeispiele Mathematik

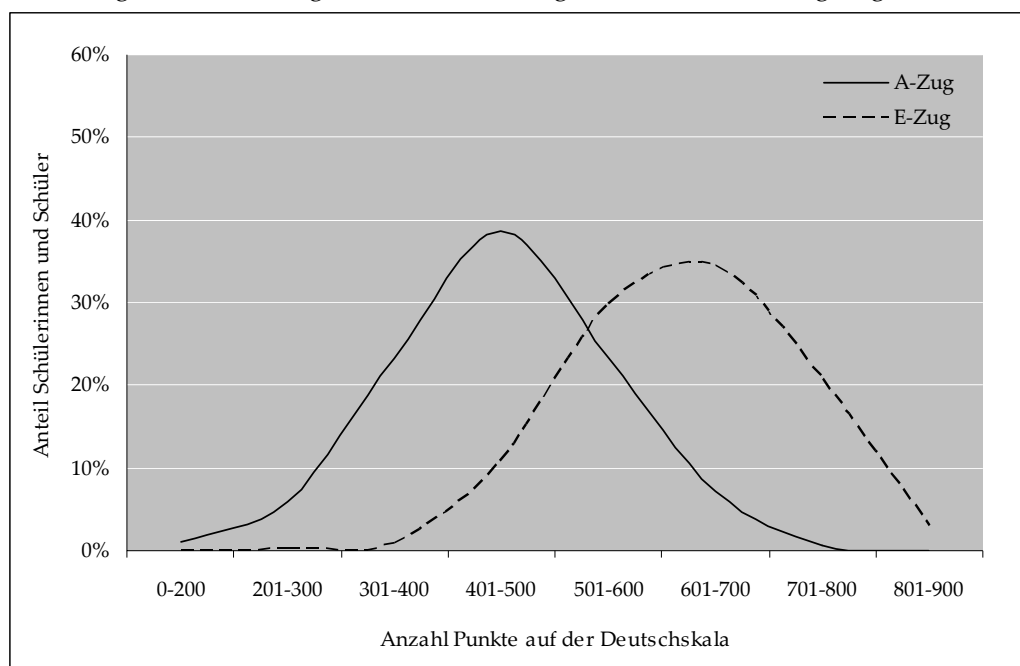
Intervall	Aufgabenschwierigkeit	Aufgabe	Anteil Schülerinnen und Schüler mit den notwendigen Fähigkeiten die Aufgabe richtig zu lösen
401 – 500 Punkte	$\delta = 414$ Punkte	<p>Stelle den gewöhnlichen Bruch als Dezimalbruch dar.</p> $\frac{3}{4} = \boxed{}$	A-Zug: 63 % E-Zug: 98 %
501 – 600 Punkte	$\delta = 549$ Punkte	<p>Stelle den gewöhnlichen Bruch in Prozenten dar.</p> $\frac{1}{20} = \boxed{}$	A-Zug: 14 % E-Zug: 68 %
601 – 700 Punkte	$\delta = 681$ Punkte	$\frac{7x}{4} - 14 = 0$	A-Zug: 1 % E-Zug: 15 %
>700 Punkte	$\delta = 731$ Punkte	<p>Berechne die Höhe h des dargestellten Dreiecks.</p>  <p>Fläche $A = 52,5 \text{ dm}^2$</p> <p>$h = ?$</p> <p>$c = 15 \text{ dm}$</p>	A-Zug: 1 % E-Zug: 6 %

Anmerkung: Die Aufgabenschwierigkeit δ zeigt die Anzahl Punkte, die eine Schülerin / ein Schüler benötigt, um die Aufgabe mit einer Wahrscheinlichkeit von 62% richtig zu lösen.

3.2 Deutschleistungen am Ende der 9. Klasse

Abbildung 3.3 zeigt die Verteilung der Deutschleistungen der Schülerinnen und Schüler in der Schlussprüfung 2011 nach Leistungszug. Die Schülerinnen und Schüler des A-Zugs erreichen am Ende der 9. Klasse durchschnittlich 451 Punkte auf der Deutschskala, die Schülerinnen und Schüler des E-Zugs 621 Punkte. Die Differenz zwischen den durchschnittlichen Deutschleistungen der Schülerinnen und Schüler des E- und des A-Zugs beträgt 170 Punkte.

Abbildung 3.3 Verteilung der Deutschleistungen 2011 nach Leistungszug

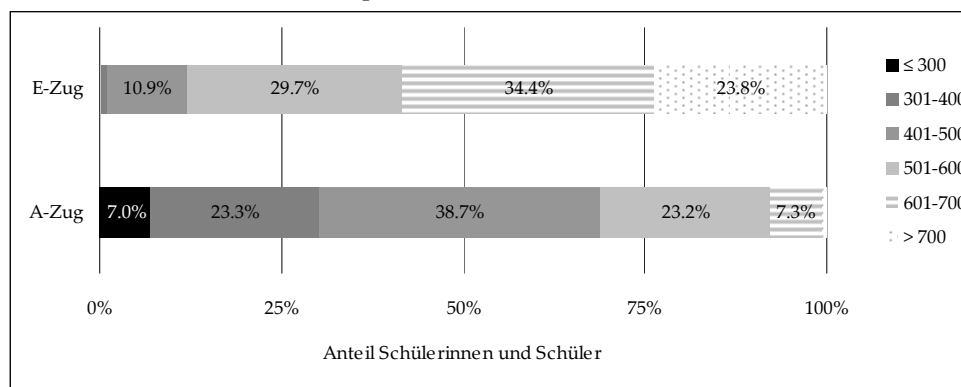


Anmerkungen: ohne Schülerinnen und Schüler der Fremdsprachenklassen
 A-Zug: N = 325; M = 451 Punkte; SD = 103 Punkte
 E-Zug: N = 456; M = 621 Punkte; SD = 99 Punkte

Die Leistungsunterschiede zwischen den Schülerinnen und Schülern des A- und des E-Zugs sind in Deutsch zwar etwas grösser als in der Mathematik. Die Leistungsverteilungen der beiden Züge überschneiden sich aber auch in Deutsch beträchtlich. Rund 40 Prozent der Schülerinnen und Schüler können aufgrund ihrer Deutschleistungen nicht eindeutig dem A- oder dem E-Zug zugewiesen werden. Rund 3 Prozent der Schülerinnen und Schüler des A-Zugs erzielen in Deutsch eine höhere Leistung als der Durchschnitt der Schülerinnen und Schüler des E-Zugs. Umgekehrt erreichen rund 4 Prozent der Schülerinnen und Schüler des E-Zugs tiefere Deutschleistungen als der Durchschnitt der Schülerinnen und Schüler des A-Zugs.

Abbildung 3.4 zeigt die Verteilung der Schülerinnen und Schüler am Ende der 9. Klasse auf die Intervalle in Deutsch nach Leistungszug. Schülerinnen und Schüler, die einem bestimmten Intervall zugeordnet werden, können die Aufgaben in diesem Intervall sowie alle Aufgaben der tieferen Intervalle mit einer Wahrscheinlichkeit von mindestens 50 Prozent richtig lösen.

Abbildung 3.4 Verteilung der Schülerinnen und Schüler auf die Intervalle in Deutsch (ohne Fremdsprachenklassen)



Rund 30 Prozent der Schülerinnen und Schüler des A-Zugs erreichen am Ende der WBS weniger als 400 Punkte auf der Deutschskala. Sie verstehen zwar den Inhalt eines einfachen Sachtextes. Allerdings fehlen ihnen in Deutsch jene Grundkompetenzen, um Fälle, Zeitformen, Pronomen oder Satzzeichen korrekt anzuwenden.

99 Prozent der Schülerinnen und Schüler des E-Zugs sowie rund 70 Prozent der Schülerinnen und Schüler des A-Zugs erreichen am Ende der WBS mehr als 400 Punkte. Sie kennen unter anderem einige Kommaeregeln sowie die wichtigsten Zeitformen und können diese in einem Text erkennen. Zudem verstehen sie einen einfachen Text so weit, dass sie zwischen den Informationen im Text Beziehungen herzustellen können.

Rund 88 Prozent der Schülerinnen und Schüler des E-Zugs sowie rund 31 Prozent der Schülerinnen und Schüler des A-Zugs erreichen am Ende der WBS mehr als 500 Punkte. Sie erkennen die inhaltliche Intention eines Textes sowie die Zeitformen von Verben. Zudem werden Fälle erkannt und richtig angewendet.

Rund 58 Prozent der Schülerinnen und Schüler des E-Zugs sowie rund 8 Prozent der Schülerinnen und Schüler des A-Zugs erreichen am Ende der WBS mehr als 600 Punkte. Sie verstehen unter anderem komplexere Texte und können Zeitformen, Fälle und Pronomen richtig anwenden.

Rund 24 Prozent der Schülerinnen und Schüler des E-Zugs erreichen am Ende der WBS mehr als 700 Punkte. Sie sind unter anderem in der Lage, in einen Text eingebettete Informationen zu finden und sprachliche Nuancen zu verstehen. Zudem können sie den Sinn schwieriger Fremdwörter aus dem kontextualen Zusammenhang herleiten.

In Tabelle 3.2 sind für jedes Intervall in Deutsch Beispiele typischer Aufgaben aus der WBS Schlussprüfung 2011 dargestellt.

Tabelle 3.2 Aufgabenbeispiele Deutsch

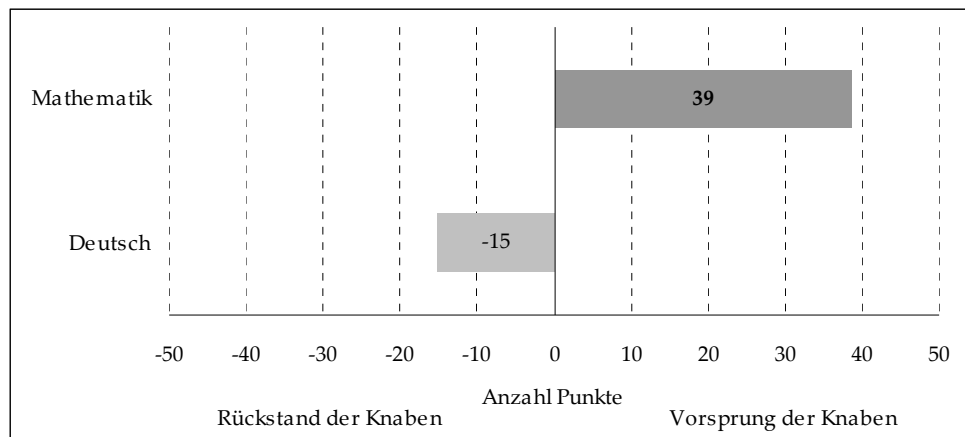
Intervall	Aufgabenschwierigkeit	Aufgabe	Anteil Schülerinnen und Schüler mit den notwendigen Fähigkeiten die Aufgabe richtig zu lösen
401 – 500 Punkte	$\delta = 421$ Punkte	<p><i>Die Verben in Klammern stehen alle in der Grundform. Schreibe eine korrekte Lösung auf die leeren Linien. Achtung: auf einer Linie können auch zwei oder mehr Wörter stehen!</i></p> <p>Bis zum heutigen Tage _____ der FCB schon 13 Mal Schweizer Meister und zehn Mal Cupsieger (werden).</p>	A-Zug: 63% E-Zug: 98%
501 – 600 Punkte	$\delta = 535$ Punkte	<p><i>Die Ausdrücke in Klammern stehen alle im Nominativ (Wer-Fall). Setze sie in den richtigen Fall und schreibe die Lösung auf die leeren Linien. Achtung: Der Nominativ selbst kommt auch vor.</i></p> <p>Dieser hatte in (viele Heimatfilme) _____ mitgespielt und war ein Liebling (ältere Damen) _____.</p>	A-Zug: 23% E-Zug: 80%
601 – 700 Punkte	$\delta = 620$ Punkte	<p>Er sollte das Porträt (ein bekannter Filmschauspieler) _____ malen.</p>	A-Zug: 6 % E-Zug: 51%
>700 Punkte	$\delta = 725$ Punkte	<p>Welches Wort oder welcher Ausdruck wurde für die folgenden Umschreibungen im Text auf Seite 4 gebraucht? Die Reihenfolge der Wörter entspricht der Reihenfolge, wie sie im Text vorkommen. Schreibe das Wort beziehungsweise den Ausdruck richtig auf die Linie!</p> <p>rentabel geplant und durchgeführt* _____</p>	A-Zug: 0% E-Zug: 16%

Anmerkung: Die Aufgabenschwierigkeit δ zeigt die Anzahl Punkte, die eine Schülerin / ein Schüler benötigt, um die Aufgabe mit einer Wahrscheinlichkeit von 62% richtig zu lösen.

3.3 Leistungsunterschiede zwischen Mädchen und Knaben am Ende der 9. Klasse

Abbildung 3.5 zeigt die Leistungsunterschiede zwischen Mädchen und Knaben am Ende der 9. Klasse als Punkte auf der WBS-Skala.

Abbildung 3.5 Leistungsunterschiede zwischen Mädchen und Knaben in der Mathematik und in Deutsch (ohne Fremdsprachenklassen)



Anmerkungen: Mathematik:
 Mädchen: N = 385; M = 516 Punkte; SD = 111 Punkte
 Knaben: N = 390; M = 555 Punkte; SD = 116 Punkte

Deutsch:
 Mädchen: N = 386; M = 558 Punkte; SD = 130 Punkte
 Knaben: N = 393; M = 543 Punkte; SD = 133 Punkte

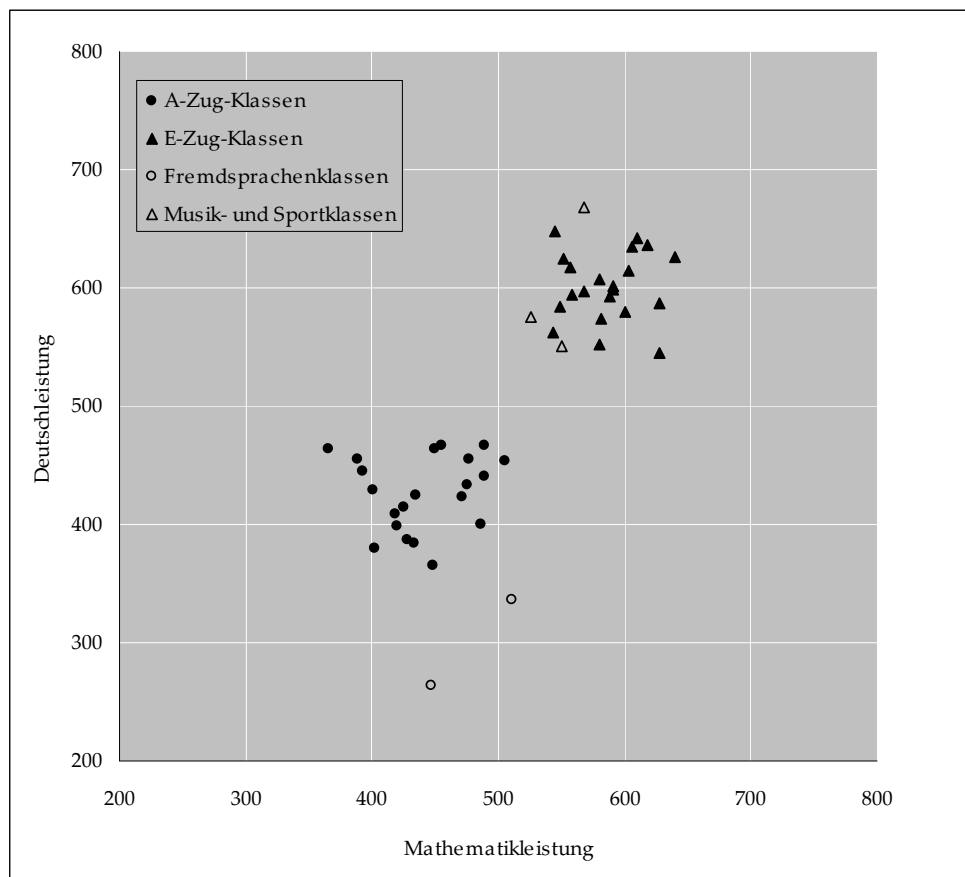
In der Mathematik erreichen die Knaben 39 Punkte mehr als die Mädchen. Dieser Unterschied ist statistisch signifikant und mit einer Effektstärke von $d = 0.34$ mittelgross. In Deutsch unterscheiden sich die Leistungen der Mädchen nicht von den Leistungen der Knaben. Der Leistungsrückstand der Knaben von 15 Punkten ist statistisch nicht signifikant.

4 Fachleistungen nach Klassen

4.1 Mathematik- und Deutschleistungen nach Klassen

Abbildung 4.1 zeigt die Ergebnisse der beteiligten Klassen in Mathematik und Deutsch. Die Ergebnisse der Regelklassen des A-Zugs werden als Punkte, die Ergebnisse der Regelklassen des E-Zugs als Dreiecke dargestellt. Jeder Punkt beziehungsweise jedes Dreieck steht für eine Klasse. Die Position der Punkte und Dreiecke ergibt sich aus der durchschnittlichen Punktzahl der Schülerinnen und Schüler einer Klasse in Deutsch und in der Mathematik.

Abbildung 4.1 Klassenmittelwerte in Deutsch und Mathematik



Anmerkungen: A-Zug-Klassen:
 Mathematik: N = 21 Klassen; M = 441 Punkte; SD = 39 Punkte
 Deutsch: N = 21 Klassen; M = 426 Punkte; SD = 32 Punkte
 E-Zug-Klassen:
 Mathematik: N = 21 Klassen; M = 587 Punkte; SD = 29 Punkte
 Deutsch: N = 21 Klassen; M = 601 Punkte; SD = 29 Punkte

Die durchschnittlichen Leistungen in den Klassen des A- und des E-Zugs unterscheiden sich stark voneinander. Kein Klassenmittelwert einer E-Klasse liegt in der Mathematik oder in Deutsch unter 500 Punkten. Hingegen wird nur in einer A-Klasse in der Mathematik eine durchschnittliche Leistung von mehr als 500 Punkten erreicht. Die Differenz

zwischen der besten A-Klasse und der schwächsten E-Klasse beträgt in der Mathematik 38 Punkte und in Deutsch 77 Punkte.

Die Klassen des A-Zugs erreichen in der Mathematik durchschnittlich 441 Punkte und in Deutsch 426 Punkte. Allerdings sind die Leistungsunterschiede zwischen den A-Klassen gross, insbesondere in der Mathematik. Die Spannweite zwischen dem höchsten und dem tiefsten Klassenmittelwert beträgt in der Mathematik 140 Punkte und in Deutsch 102 Punkte.

Die Klassen des E-Zugs erreichen in der Mathematik im Durchschnitt 587 Punkte und in Deutsch 601 Punkte. Die Leistungen der E-Klassen sind einheitlicher als die Leistungen der A-Klassen. D.h. die Streuung zwischen den Klassen (Standardabweichung) des E-Zugs ist in der Mathematik deutlich geringer als zwischen den Klassen des A-Zugs. Die Spannweite zwischen dem höchsten und dem tiefsten Klassenmittelwert des E-Zugs beträgt in der Mathematik 97 Punkte. In Deutsch hingegen ist die Differenz zwischen dem höchsten und dem tiefsten Klassenmittelwert mit 104 Punkten etwa gleich gross wie im A-Zug.

Die Leistungen der beiden Fremdsprachenklassen sind in Deutsch klar schwächer als die Deutschleistungen der A-Klassen. In der Mathematik hingegen unterscheiden sich die Leistungen der Fremdsprachenklassen nicht von den Leistungen der A-Klassen. Eine Fremdsprachenklasse erreicht in der Mathematik durchschnittlich 511 Punkte und liegt damit über dem Klassenmittelwert der besten A-Klasse.

Die Leistungen der Musik- und Sportklassen mit den Leistungen der Regelklassen zu vergleichen, ist schwierig, da in den Musik- und Sportklassen sowohl Schülerinnen und Schüler des A-Zugs als auch Schülerinnen und Schüler des E-Zugs unterrichtet werden. Zudem unterscheidet sich der Anteil Mädchen und Knaben teilweise stark von jenem in den Regelklassen. Insgesamt weichen die Leistungen der Musik- und Sportklassen jedoch kaum von den Leistungen der E-Klassen ab. In der Mathematik (Mittelwert $M = 548$ Punkte) zählen die Musik- und Sportklassen eher zu den schwächeren Klassen des E-Zugs. In Deutsch liegen die Klassenmittelwerte zweier Musik- und Sportklassen unter dem Mittelwert aller E-Klassen. Eine der Musik- und Sportklassen jedoch erreicht mit rund 668 Punkten das beste Deutschergebnis aller Klassen der WBS.

4.2 Beurteilung der Unterschiede zwischen den Klassen

Die unterschiedlichen Leistungen der Klassen lassen sich auf Merkmale der Schülerinnen und Schüler und auf Merkmale der Klassen zurückführen. Je stärker die Leistungen zwischen den Klassen variieren, desto eher können sie durch Merkmale der Klasse beziehungsweise durch Merkmale des Unterrichts erklärt werden.

Tabelle 4.1 Varianzen zwischen und innerhalb der Klassen:
Mathematik und Deutsch nach Leistungszug

	Varianz zwischen den Klassen	Varianz zwischen den Schülerinnen und Schülern
E-Zug Mathematik	6%	94%
E-Zug Deutsch	3%	97%
A-Zug Mathematik	13%	87%
A-Zug Deutsch	3%	97%

Tabelle 4.1 zeigt für Mathematik und für Deutsch die Prozentanteile, die entweder durch individuelle Merkmale der Schülerinnen und Schüler oder durch Unterrichts- und Klassenmerkmale erklärt werden können.

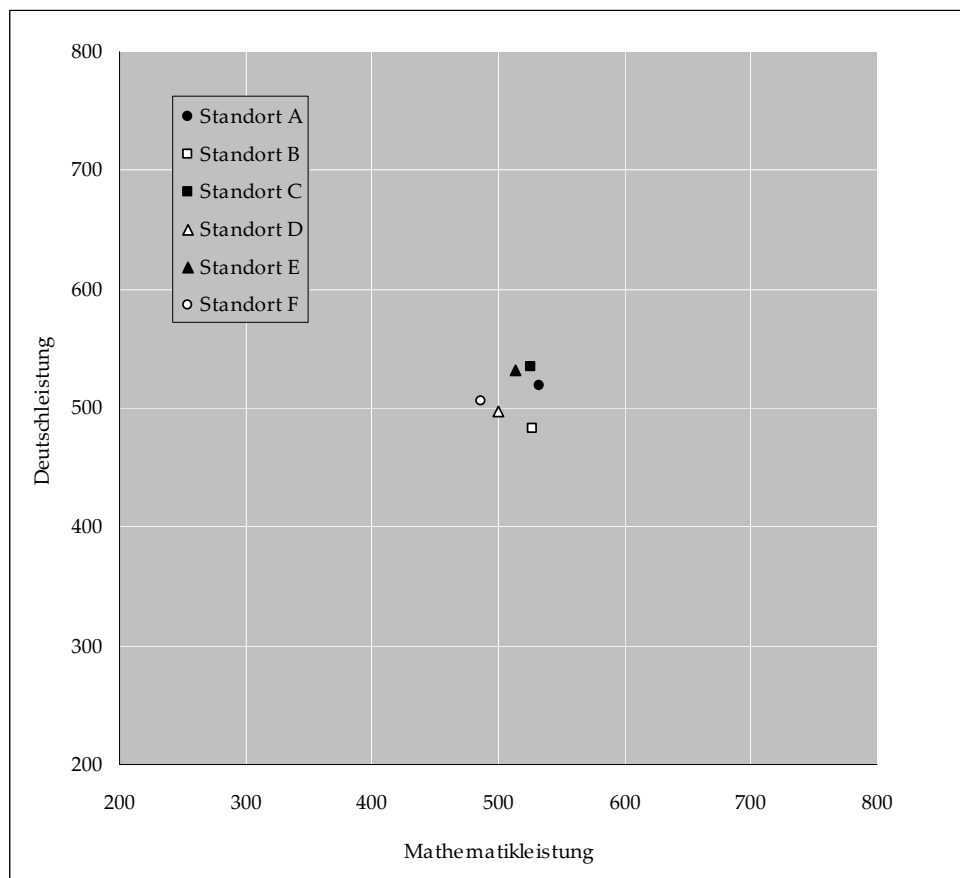
Im E-Zug können in der Mathematik 6 Prozent der Leistungsunterschiede auf Merkmale des Unterrichts oder der Klasse zurückgeführt werden. 94 Prozent können mit individuellen Merkmalen der Schülerinnen und Schüler erklärt werden. In Deutsch können im E-Zug 3 Prozent der Leistungsunterschiede mit Merkmalen des Unterrichts und 97 Prozent mit Merkmalen der Schülerinnen und Schüler erklärt werden.

Im A-Zug können in der Mathematik 13 Prozent der Leistungsunterschiede mit Merkmalen des Unterrichts oder der Klasse erklärt werden. Das ist ein deutlich grösserer Anteil als in Deutsch und ein doppelt so grosser Anteil als zwischen den Klassen des E-Zugs. Das heisst, dass (1) im A-Zug die Merkmale des Unterrichts oder der Klasse wichtiger sind, um die Leistungsunterschiede in der Mathematik zu erklären als im E-Zug, dass es (2) für den Erfolg in den Schlussprüfungen relevant ist, welche Klasse die Schülerinnen und Schüler des A-Zugs besuchen und dass (3) vermutlich Potenzial besteht, um mit geeigneten Unterrichtsmassnahmen die Mathematikleistungen der Schülerinnen und Schüler des A-Zugs zu steigern.

5 Fachleistungen nach WBS-Standorten

Abbildung 5.1 zeigt die Ergebnisse der Schülerinnen und Schüler der Regelklassen pro WBS-Standort. Die Position eines Standorts ergibt sich aus der durchschnittlichen Punktzahl, die an einem Standort in Mathematik und Deutsch erreicht wurde. Um den Einfluss der unterschiedlichen grossen Anteile an A- und E-Klassen auf das Ergebnis eines Standorts auszugleichen, wurde das Leistungsniveau der Schülerinnen und Schüler statistisch kontrolliert. Das heisst, die Ergebnisse von Standorten mit einem hohen Anteil Schülerinnen und Schüler des A-Zugs wurden nach oben korrigiert, die Ergebnisse von Standorten mit einem tiefen Anteil Schülerinnen und Schüler des A-Zugs wurden entsprechend nach unten korrigiert.

Abbildung 5.1 Fachleistungen in Deutsch und Mathematik nach WBS-Standorten (nach statistischer Kontrolle der Leistungsniveaus)



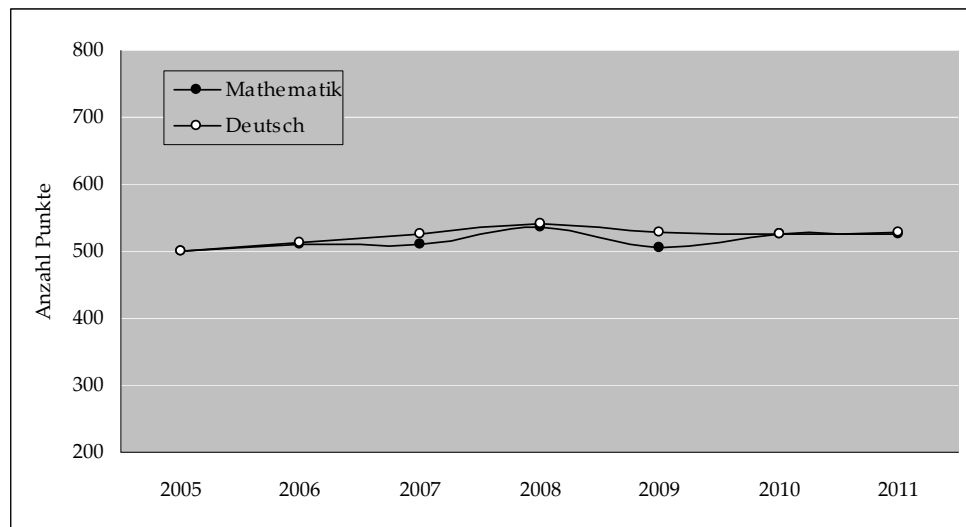
In der Mathematik sind die Mathematikleistungen am Standort F am tiefsten. Sie unterscheiden sich statistisch signifikant von den Mathematikleistungen an den Standorten A, B und C, jenen drei Standorten mit den besten Leistungen in Mathematik. Die Differenz zwischen dem höchsten und dem tiefsten Mittelwert beträgt 46 Punkte. In Deutsch werden an den Standorten C und E statistisch signifikant bessere Leistungen erbracht, als an den Standorten B, D und F. Die Deutschleistungen am Standort B liegen rund 24 Punkte unter dem Gesamtmittelwert aller WBS-Standorte und rund 51 Punkte unter den durchschnittlichen Leistungen des Standorts C, dem Standort mit den höchsten Deutschleistungen ($M = 535$ Punkte).

6 Leistungsentwicklung an der WBS

6.1 Entwicklung der durchschnittlichen Leistungen der Schülerinnen und Schüler an der WBS

Abbildung 6.1 zeigt die durchschnittlichen Leistungen der Schülerinnen und Schüler in den Fächern Mathematik und Deutsch in den Jahren 2005 bis 2011. Die Mittelwerte in Mathematik sind als schwarze Punkte, die Mittelwerte in Deutsch als weisse Punkte dargestellt.

Abbildung 6.1 Leistungsentwicklung in den Fächern Mathematik und Deutsch



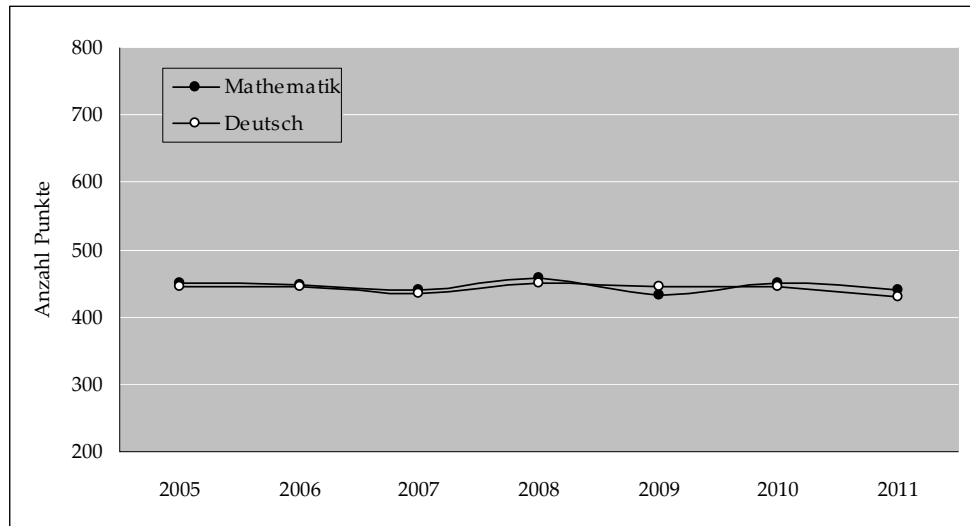
Anmerkung: Die Daten zu dieser Abbildung finden sich im Anhang

In der Mathematik haben sich die durchschnittlichen Leistungen seit 2005 um insgesamt 24 Punkte auf 524 Punkte erhöht. Diese Leistungssteigerung ist statistisch signifikant und mit einer Effektgrösse von $d = 0.25$ zwar gering, aber durchaus von Relevanz. Die Leistungssteigerung verlief aber nicht linear, sondern war Schwankungen zwischen den einzelnen Prüfungsjahren unterlegen. Auffällig sind insbesondere die sehr guten Mathematikleistungen im Jahr 2008 (535 Punkte). Zwischen 2010 und 2011 sanken die durchschnittlichen Mathematikleistungen an der WBS um einen Punkt auf durchschnittlich 524 Punkte. Diese Leistungsveränderung ist statistisch nicht signifikant.

In Deutsch erhöhten sich die durchschnittlichen Leistungen seit 2005 um 29 Punkte auf 529 Punkte (Effektgrösse $d = 0.25$). Auch in Deutsch verlief die Leistungsentwicklung nicht linear. Bis zum Jahr 2008 stiegen die Leistungen in Deutsch relativ kontinuierlich an. Danach sanken sie wieder. Seit 2009 bleiben die Deutschleistungen an der WBS weitgehend konstant.

Abbildung 6.2 zeigt die durchschnittlichen Leistungen der Schülerinnen und Schüler des A-Zugs in Mathematik und Deutsch in den Jahren 2005 bis 2011.

Abbildung 6.2 Leistungsentwicklung in den Fächern Mathematik und Deutsch im A-Zug

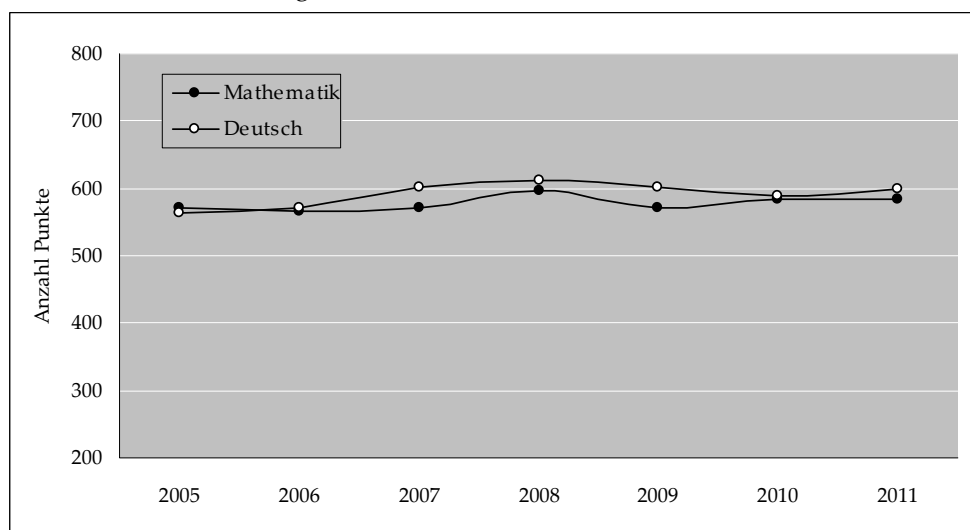


Anmerkung: Die Daten zu dieser Abbildung finden sich im Anhang

Im A-Zug sind die Mathematikleistungen 2011 gegenüber 2010 um 12 Punkte auf 439 Punkte gesunken. Sie unterscheiden sich damit nicht statistisch signifikant von den Mathematikleistungen 2009, als an der WBS die Massnahmen zur Steigerung der Mathematikleistungen erstmals umgesetzt wurden. In Deutsch sind die Leistungen um 16 Punkte gesunken und sind damit so tief wie noch nie seit 2005.

In Abbildung 6.3 sind die durchschnittlichen Leistungen der Schülerinnen und Schüler des E-Zugs in den Jahren 2005 bis 2011 dargestellt.

Abbildung 6.3 Leistungsentwicklung in den Fächern Mathematik und Deutsch im E-Zug



Anmerkung: Die Daten zu dieser Abbildung finden sich im Anhang

Im E-Zug haben sich die Mathematikleistungen gegenüber 2010 nicht verändert. In Deutsch stiegen die Leistungen des E-Zugs gegenüber dem Vorjahr um 10 Punkte. Damit sind die durchschnittlichen Deutschleistungen im E-Zug wieder gleich hoch wie in den Jahren 2007 und 2009.

6.2 Entwicklung der Leistungsunterschiede zwischen Mädchen und Knaben

Abbildung 6.4 zeigt die Leistungsunterschiede zwischen Mädchen und Knaben in der Mathematik und in Deutsch seit 2005. Die dunkelgrauen Balken zeigen den durchschnittlichen Leistungsvorsprung der Knaben in der Mathematik als Anzahl Punkte. Die hellgrauen Balken symbolisieren den durchschnittlichen Leistungsrückstand der Knaben in Deutsch.

Abbildung 6.4 Leistungsunterschiede zwischen Mädchen und Knaben in der Mathematik und in Deutsch 2005 bis 2011

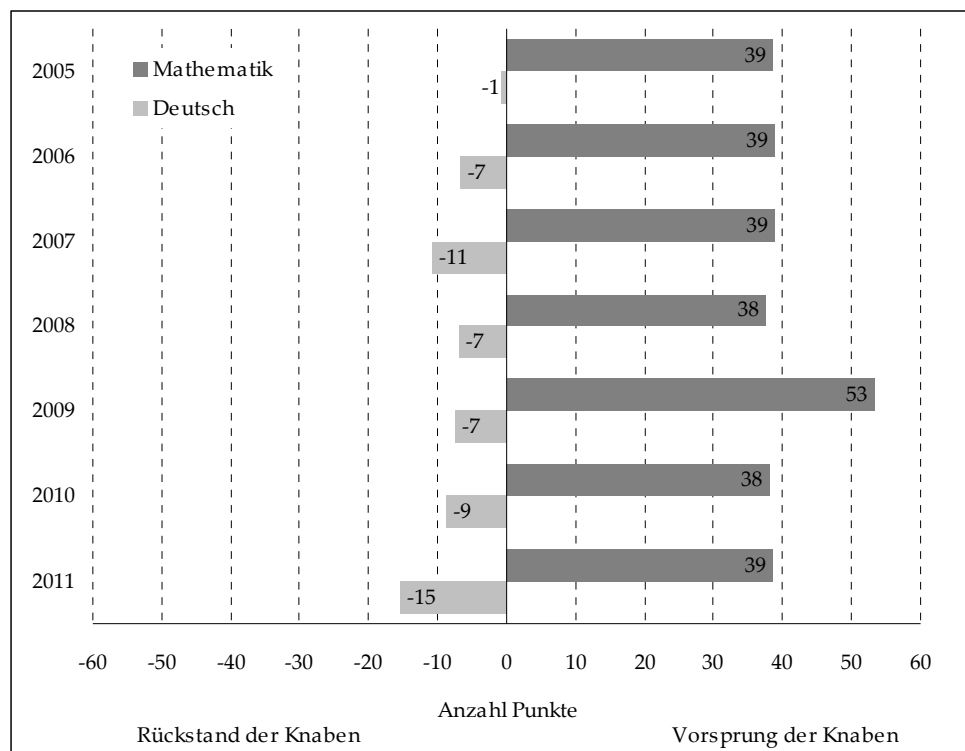


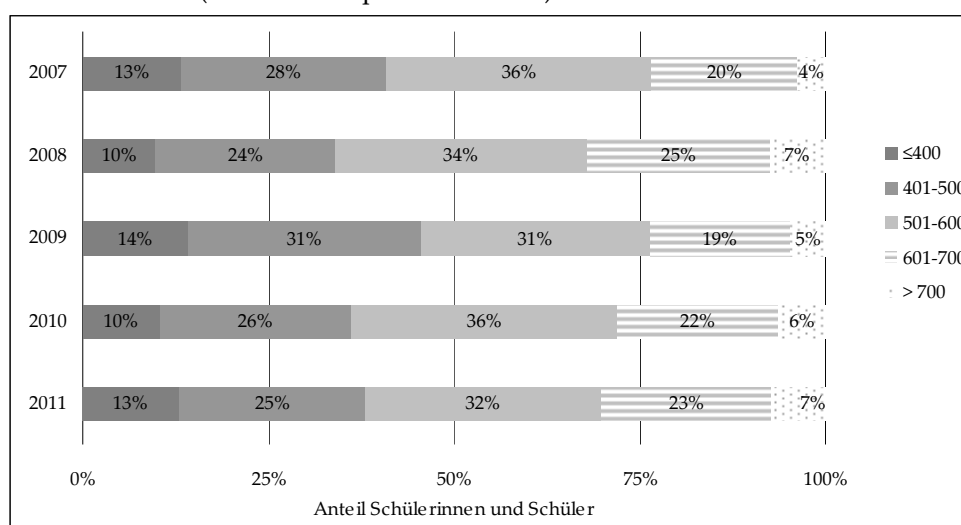
Abbildung 6.4 zeigt zum einen, dass die Unterschiede zwischen den Geschlechtern seit 2005 weitgehend konstant sind. In der Mathematik sind die Knaben im Durchschnitt jeweils 38 bis 39 Punkte besser als die Mädchen. Einzig im Jahr 2009 betrug der Leistungsvorsprung der Knaben 53 Punkte. In Deutsch variiert der Leistungsrückstand der Knaben zwischen 1 und 15 Punkten.

Zum andern sind die Leistungsunterschiede zwischen Mädchen und Knaben in der Mathematik viel grösser als in Deutsch. Der Leistungsvorsprung der Knaben in der Mathematik ist bei einer Effektgrösse von $d = 0.34$ (2011) bis $d = 0.50$ (2009) statistisch signifikant. In Deutsch sind die Leistungsunterschiede zwischen den Geschlechtern hingegen in keinem Jahr statistisch signifikant.

6.3 Entwicklung der Fähigkeiten der Schülerinnen und Schüler an der WBS

Neben der Entwicklung der durchschnittlichen Leistungen ist auch von Interesse, ob sich der Anteil Schülerinnen und Schüler, der in der Mathematik oder in Deutsch nicht über die notwendigen Grundkompetenzen verfügt, vergrössert oder verringert. Um diese Frage zu beantworten, wurden die Schülerinnen und Schüler aufgrund ihrer Leistungen in Intervalle von jeweils 100 Punkten eingeteilt. Abbildung 6.5 zeigt die Verteilung der Schülerinnen und Schüler in der Mathematik nach Intervall in den Jahren 2007 bis 2011.

Abbildung 6.5 Verteilung der Schülerinnen und Schüler nach Intervall in der Mathematik 2007 bis 2011 (ohne Fremdsprachenklassen)

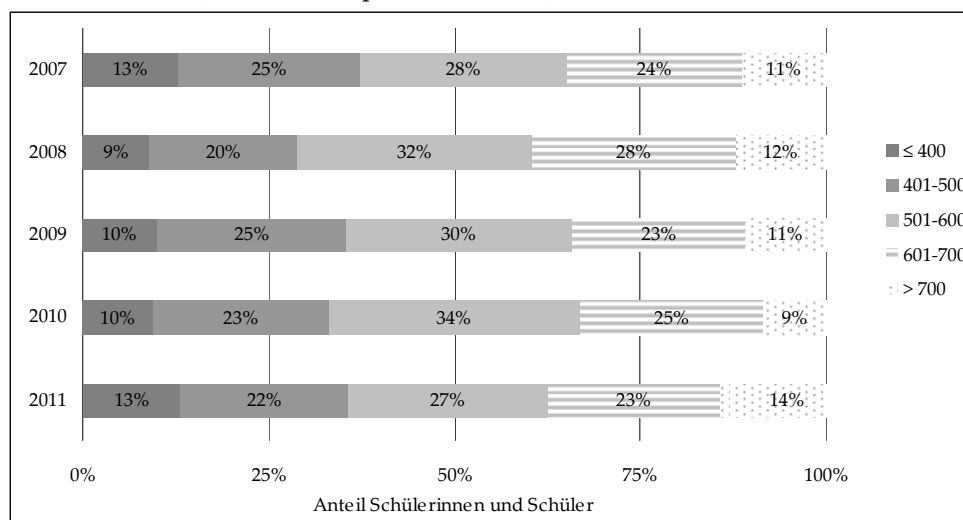


Schülerinnen und Schüler, die in der Mathematik 400 oder weniger Punkte erreichen, verfügen nicht über mathematische Grundkompetenzen. Diese Schülerinnen und Schüler sind am Ende der obligatorischen Schule nicht in der Lage, einfache Grundoperationen zu lösen. 2011 erreichen 13 Prozent der Schülerinnen und Schüler weniger als 400 Punkte. Gegenüber dem Vorjahr ist der Anteil an Schülerinnen und Schüler mit ungenügenden Grundkompetenzen um 3 Prozent gestiegen.

Die besten Schülerinnen und Schüler erreichen in der Mathematik mehr als 600 Punkte. Sie sind in der Lage, Bruchgleichungen zu lösen und Folgerungen aus grafischen Darstellungen statistischer Daten zu ziehen. Der Anteil Schülerinnen und Schüler, die mehr als 600 Punkte erreichen, stieg von 28 Prozent im Jahr 2010 auf 30 Prozent im Jahr 2011.

Abbildung 6.6 zeigt die Verteilung der Schülerinnen und Schüler in Deutsch nach Intervall in den Jahren 2007 bis 2011.

Abbildung 6.6 Verteilung der Schülerinnen und Schüler nach Intervall in Deutsch 2007 bis 2011 (ohne Fremdsprachenklassen)



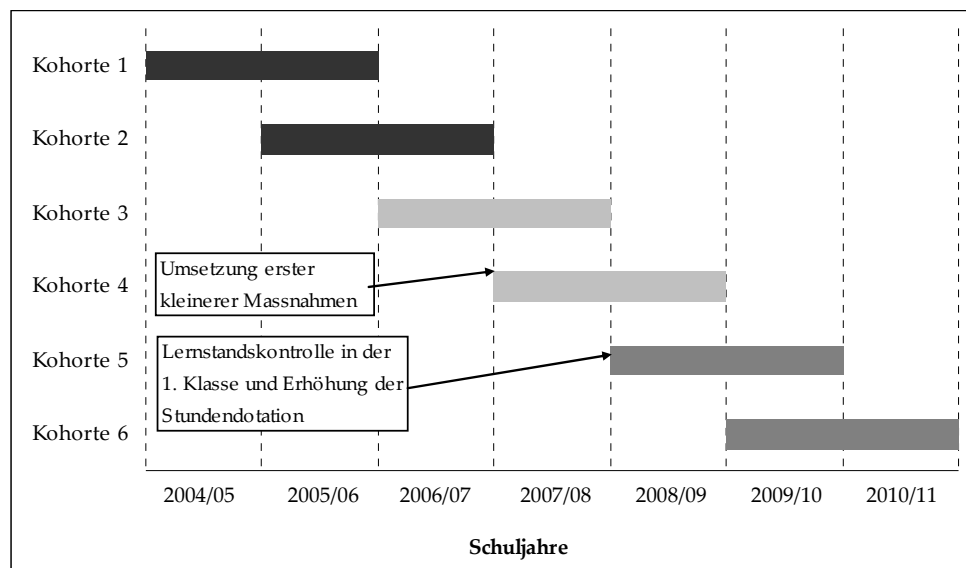
Schülerinnen und Schüler mit 400 oder weniger Punkten verfügen am Ende der obligatorischen Schulzeit nicht über die notwendigen Grundkompetenzen in Deutsch. Sie sind nicht in der Lage, einen längeren Text zu verstehen sowie grundlegende grammatikalische und syntaktische Regeln anzuwenden, wie beispielsweise Substantive in den korrekten Fall oder Verben in die korrekte Zeitform zu setzen. 2011 erreichen 13 Prozent der Schülerinnen und Schüler in Deutsch weniger als 400 Punkte. Dieser Anteil ist rund 3 Prozent grösser als 2010.

Die besten Schülerinnen und Schüler erreichen in Deutsch mehr als 600 Punkte. Sie verstehen komplexere Texte und können Zeitformen, Fälle und Pronomen richtig anwenden. Ihr Anteil betrug im Jahr 2011 37 Prozent. Damit ist auch der Anteil Schülerinnen und Schüler mit mehr als 600 Punkten gegenüber dem Vorjahr um 3 Prozent angestiegen.

7 Leistungsentwicklung an der WBS seit der Umsetzung des «Mathe-Pakets»

Aufgrund der ungenügenden Mathematikleistungen am Ende der 9. Klasse ergriff das Rektorat der WBS 2007 unter dem Namen «Mathe-Paket» mehrere Massnahmen, um den Mathematikunterricht zu stärken und damit die Mathematikleistungen nachhaltig zu steigern. Kleinere Massnahmen wie die Formulierung klarer Leistungserwartungen für die Schlussprüfungen durch die Schulleitungen sowie die Kommunikation der in den Schlussprüfungen geforderten Fähigkeiten wurden bereits im Schuljahr 2007/08 umgesetzt. Die einschneidendsten Massnahmen des «Mathe-Pakets», wie die obligatorische Durchführung einer Lernstandserhebung zu Beginn der 1. Klasse der WBS als Grundlage für einen förderorientierten Unterricht, das Zusammenstellen von Aufgabensammlungen als Übungsmaterial für die Schlussprüfungen und die Erhöhung der Stundendotation in der Mathematik um eine Lektion wurden auf das Schuljahr 2008/09 eingeführt. Das heisst, die Jugendlichen, die im Juni 2011 die Schlussprüfungen an der WBS lösten, sind die zweite Schülerkohorte, die von den Massnahmen des «Mathe-Pakets» vollumfänglich profitieren konnte (Abbildung 7.1).

Abbildung 7.1 Umsetzung des «Mathe-Pakets» seit 2007 und Einfluss auf die Schülerkohorten



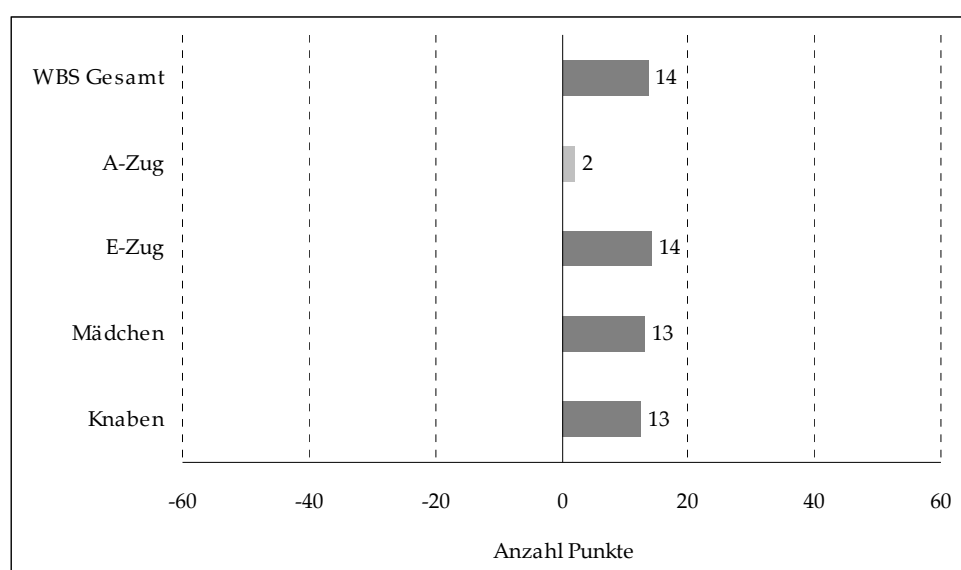
Anmerkung: Kohorte 1 = Schlussprüfungen 2006 (WBS ohne «Mathe-Paket»)
 Kohorte 2 = Schlussprüfungen 2007 (WBS ohne «Mathe-Paket»)
 Kohorte 3 = Schlussprüfungen 2008 (WBS z.T. mit «Mathe-Paket»)
 Kohorte 4 = Schlussprüfungen 2009 (WBS z.T. mit «Mathe-Paket»)
 Kohorte 5 = Schlussprüfungen 2010 (WBS mit «Mathe-Paket»)
 Kohorte 6 = Schlussprüfungen 2011 (WBS mit «Mathe-Paket»)

Um die Wirksamkeit der getroffenen Massnahmen zu überprüfen, werden die Leistungen der Kohorten 1 und 2, die die Schlussprüfungen im Juni 2006 bzw. im Juni 2007 abgelegt haben, mit den Leistungen der Kohorten 5 und 6 verglichen. Die Schülerinnen und Schüler der Kohorten 1 und 2 durchliefen die WBS ohne zusätzliche Massnahmen zur Steigerung der Mathematikleistungen. Die Schülerinnen und Schüler der Kohorten 5 und 6 hingegen konnten während ihrer gesamten Schulzeit an der WBS von den Mas-

snahmen des «Mathe-Pakets» profitieren. Vorausgesetzt die Massnahmen zur Steigerung der Mathematikleistung haben die intendierten Wirkungen, dann müssten die Leistungen der beiden Kohorten *nach* Umsetzung des «Mathe-Pakets» bessere Leistungen erbringen als die beiden Kohorten *vor* der Umsetzung.

Abbildung 7.2 zeigt die Leistungsveränderungen in der Mathematik zwischen den Schülerinnen und Schülern, die vor der Umsetzung des «Mathe-Pakets» die WBS abschlossen (Kohorten 1 und 2) und den Schülerinnen und Schüler, die erst nach der Umsetzung des «Mathe-Pakets» in die WBS eintraten (Kohorten 5 und 6).

Abbildung 7.2 Veränderung der Mathematikleistungen vor und nach der Umsetzung des «Mathe-Pakets»

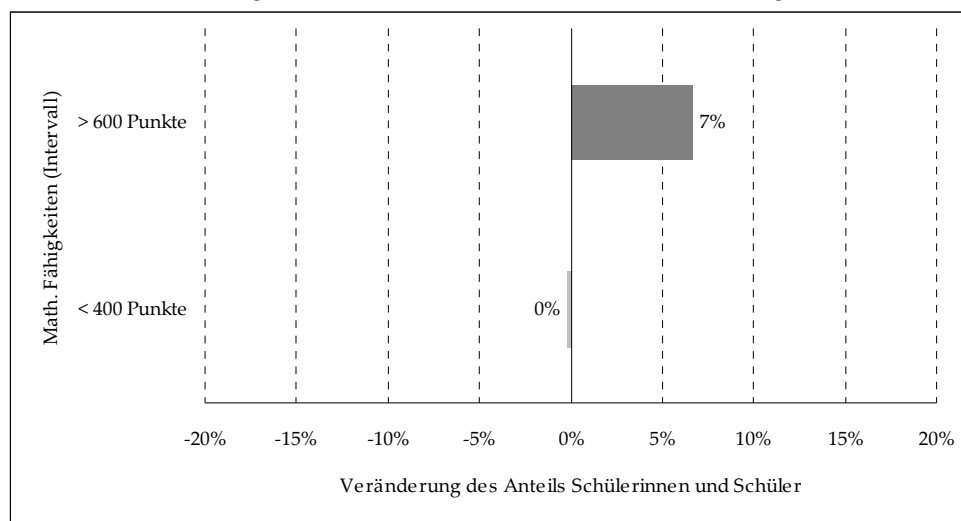


Anmerkung: dunkelgraue Balken = statistisch signifikante Veränderungen
hellgraue Balken = statistisch nicht signifikante Veränderungen

Wie Abbildung 7.2 zeigt, haben sich die Mathematikleistungen an der WBS seit der Einführung des «Mathe-Pakets» um 14 Punkte verbessert. Diese Leistungssteigerung ist statistisch signifikant, mit einer Effektgrösse von $d = 0.13$ jedoch gering. Interessant ist, dass sich die Leistungssteigerung nicht in beiden Leistungszügen gleichermassen nachweisen lässt. Während die Schülerinnen und Schüler des E-Zugs nach der Umsetzung des «Mathe-Pakets» rund 14 Punkte besser sind als vor der Umsetzung, verändern sich die Mathematikleistungen der Schülerinnen und Schüler des A-Zugs nicht¹. Das könnte ein Hinweis sein, dass vor allem gute Schülerinnen und Schüler von den Massnahmen des «Mathe-Pakets» profitieren. Darauf deuten auch die Veränderungen des Anteils Schülerinnen und Schüler mit sehr schwachen und sehr guten Fähigkeiten in der Mathematik hin. Abbildung 7.3 zeigt, um wie viel Prozent sich der Anteil jener Schülerinnen und Schüler entwickelt hat, die in der Mathematik weniger als 400 Punkte beziehungsweise mehr als 600 Punkte erreichen.

¹ Die in Abbildung 6.8 ausgewiesene Leistungssteigerung von 2 Punkten ist statistisch nicht signifikant und mit einer Effektgrösse von $d = 0.02$ praktisch nicht bedeutsam.

Abbildung 7.3 Veränderung des Anteils Schülerinnen und Schüler ausgewählter Fähigkeitsintervalle vor und nach der Umsetzung des «Mathe-Pakets»

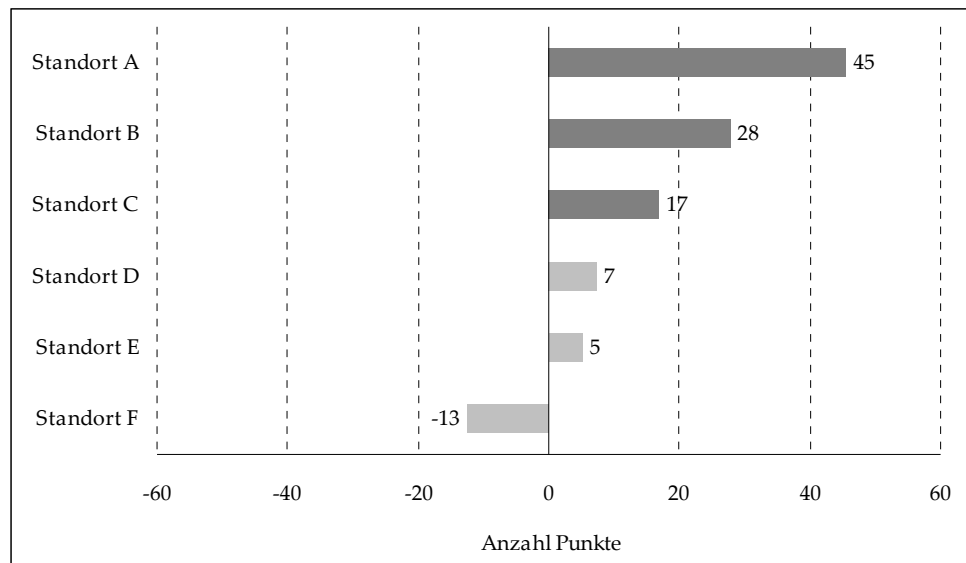


Anmerkung: dunkelgraue Balken = statistisch signifikante Veränderungen
hellgraue Balken = statistisch nicht signifikante Veränderungen

Der Anteil Schülerinnen und Schüler mit mehr als 600 Punkten in der Mathematik stieg um rund 7 Prozent. Das heisst, dass nach der Umsetzung der Massnahmen zur Steigerung der Mathematikleistungen 7 Prozent mehr Schülerinnen und Schüler sehr gute mathematische Kompetenzen haben. Der Anteil Schülerinnen und Schüler hingegen, die am Ende der WBS nicht über die notwendigen mathematischen Grundkompetenzen verfügen, hat sich trotz des «Mathe-Pakets» nicht verändert.

Die Mathematikleistungen haben sich nicht nur je nach Leistungszug, beziehungsweise je nach Fähigkeitsniveau der Schülerinnen und Schüler unterschiedlich entwickelt. Auch an den WBS-Standorten zeigen sich zwischen den Mathematikleistungen vor und nach der Umsetzung des «Mathe-Pakets» unterschiedliche Entwicklungen (Abbildung 7.4). An drei der sechs WBS-Standorten haben sich die Mathematikleistungen seit der Umsetzung des «Mathe-Pakets» statistisch signifikant verbessert. An den anderen drei Standorten kam es hingegen zu keinen signifikanten Änderungen.

Abbildung 7.4 Veränderung der Mathematikleistungen vor und nach der Umsetzung des «Mathe-Pakets» nach WBS-Standorten



Anmerkung: dunkelgraue Balken = statistisch signifikante Veränderungen
hellgraue Balken = statistisch nicht signifikante Veränderungen

Alles in allem haben sich die Mathematikleistungen an der WBS seit der Umsetzung des «Mathe-Pakets» verbessert. Die Leistungssteigerung ist zwar klein, aber statistisch signifikant. Einschränkend muss hinzugefügt werden, dass sich die Mathematikleistungen nicht bei allen Gruppen von Schülerinnen und Schülern verbessert haben. Vor allem gute Mathematikschülerinnen und -schüler sowie Schülerinnen und Schüler an bestimmten Standorten konnten von dem Massnahmenpaket profitieren.

Diese letztlich eher enttäuschende Einschätzung der Wirksamkeit des «Mathe-Pakets» muss aber relativiert werden. Denn wie in den Berichten zu den Schlussprüfungen 2010 und 2009 ausführlich beschrieben wurde, werden die Mathematikleistungen auch von demografischen und schulpolitischen Faktoren beeinflusst, die ausserhalb des Einflussbereichs der WBS und ihrer Lehrpersonen liegen².

Insbesondere drei Faktoren haben einen relevanten Einfluss auf die Leistungen an der WBS:

Ein erster Faktor ist der Anteil Knaben an der WBS. Da Knaben in der Mathematik bessere Leistungen erzielen, führt ein höherer Anteil Knaben in einem Prüfungsjahrgang unweigerlich zu besseren Durchschnittleistungen in der Mathematik.

Der zweite externe Einflussfaktor ist der Anteil Schülerinnen und Schüler im E-Zug. Zwischen den einzelnen Jahrgängen der Schülerinnen und Schüler gibt es naturgemäss kleinere Leistungsunterschiede. Das zeigt sich u.a. darin, dass in leistungsstarken Jahrgängen ein grösserer Anteil der Schülerinnen und Schüler in die Klassen des E-Zugs

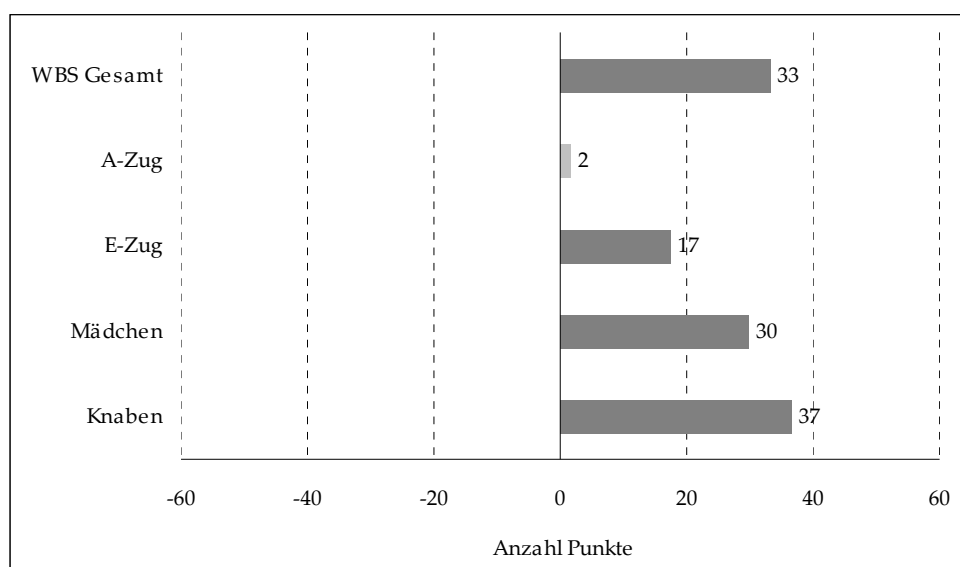
² vgl. Keller, F. & Moser, U. (2010). *Entwicklung und Durchführung der Schlussprüfungen an der Weiterbildungsschule Basel-Stadt. Schlussbericht 2010 zuhanden der Stufenleitung Sekundarstufe I des Kantons Basel-Stadt*. Zürich: Institut für Bildungsevaluation. [http://www.ibe.uzh.ch/publikationen/WBS_Schlussbericht_2010.pdf]

eingeteilt wird. Zudem sind die Lernfortschritte in den Klassen des E-Zugs grösser als in den Klassen des A-Zugs³. Dies führt dazu, dass Prüfungsjahrgänge mit einem hohen Anteil an Schülerinnen und Schülern im E-Zug bessere Durchschnittsleistungen in der Mathematik erbringen.

Der dritte und, weil grösseren Schwankungen unterworfen, wichtigste Faktor ist der Anteil Schülerinnen und Schüler, die nach der Orientierungsschule (OS) an die WBS übertreten. Geht ein grosser Anteil der Kinder nach der OS an ein Gymnasium, so besuchen tendenziell weniger leistungsstarke Schülerinnen und Schüler die WBS, was die Durchschnittsleistungen an der WBS sinken lässt. Umgekehrt führt ein hoher Anteil an Kindern in heilpädagogischen Spezialangeboten und Kleinklassen dazu, dass die Durchschnittsleistungen an der WBS steigen.

Für eine realistische Beurteilung der Wirkungen des «Mathe-Pakets» müssen diese externen Einflussfaktoren der Mathematikleistung kontrolliert werden. Die entscheidende Frage ist, wie sich die Mathematikleistungen an der WBS entwickelt hätten, wenn der Anteil Knaben, der Anteil Schülerinnen und Schüler im E-Zug und der Anteil Gymnasiastinnen und Gymnasiasten im Kanton Basel-Stadt konstant geblieben wären. Um diese Frage zu beantworten, wurde mit einer Mehrebenenanalyse die Veränderung in den Mathematikleistungen der Schülerinnen und Schüler vor und nach der Umsetzung des «Mathe-Pakets» unter Kontrolle der externen Faktoren berechnet (Abbildung 7.5).

Abbildung 7.5 Veränderung der Mathematikleistungen vor und nach der Umsetzung des «Mathe-Pakets» nach statistischer Kontrolle externer Einflussfaktoren



Anmerkung: Kontrolliert wurden der Anteil Knaben pro Testjahr, der Anteil Schülerinnen und Schüler pro Leistungszug sowie die Übertrittsquote an die WBS
dunkelgraue Balken = statistisch signifikante Veränderungen
hellgraue Balken = statistisch nicht signifikante Veränderungen

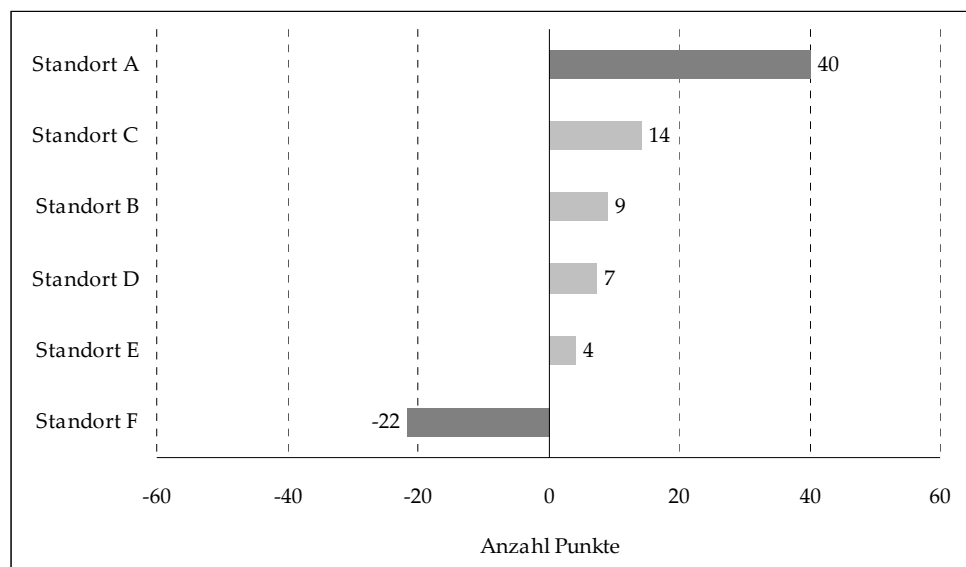
³ vgl. Moser, U. & Keller, F. (2006). *Evaluation der Strukturänderung an der Weiterbildungsschule Basel-Stadt. Schlussbericht zuhanden des Erziehungsdepartements des Kantons Basel-Stadt, Ressort Schulen*. Universität Zürich: Kompetenzzentrum für Bildungsevaluation und Leistungsmessung. [<http://www.ibe.uzh.ch/publikationen/Evaluation-WBS.pdf>]

Wie Abbildung 7.5 zeigt, hätten sich die Mathematikleistungen der Schülerinnen und Schüler an der WBS seit der Einführung des «Mathe-Pakets» in einem viel grösseren Umfang verbessert, wenn sich die soziodemografischen und schulpolitischen Voraussetzungen nicht verändert hätten. Werden die Wirkungen der externen Einflussfaktoren ausgeschlossen, beträgt der Leistungszuwachs aufgrund des «Mathe-Pakets» 33 Punkte. Diese Leistungssteigerung ist statistisch signifikant und mit einer Effektgrösse von $d = 0.32$ relevant.

Allerdings haben sich die Mathematikleistungen nicht bei allen Schülerinnen und Schülern gleichermassen verbessert. Die Mathematikleistungen der Schülerinnen und Schüler des A-Zugs haben sich mit der Umsetzung des «Mathe-Pakets» auch unter Kontrolle der externen Faktoren statistisch nicht verändert. Die Schülerinnen und Schüler des E-Zugs hingegen konnten ihre Leistungen um 17 Punkte steigern⁴. Auch zwischen den Geschlechtern zeigen sich Unterschiede. Die Mädchen konnten ihre Leistungen um durchschnittlich 30 Punkte, die Knaben um durchschnittlich 37 Punkte steigern.

Ebenso haben sich die Leistungen je nach WBS-Standort unterschiedlich entwickelt. Abbildung 7.6 zeigt die Veränderung der Mathematikleistung an den einzelnen WBS-Standorten vor und nach der Umsetzung des «Mathe-Pakets» nach statistischer Kontrolle dieser externen Faktoren.

Abbildung 7.6 Veränderung der Mathematikleistungen an den WBS-Standorten vor und nach der Umsetzung des «Mathe-Pakets» nach statistischer Kontrolle externer Einflussfaktoren



Anmerkung: Kontrolliert wurden der Anteil Knaben pro Testjahr, der Anteil Schülerinnen und Schüler pro Leistungszug sowie die Übertrittsquote an die WBS
dunkelgraue Balken = statistisch signifikante Veränderungen
hellgraue Balken = statistisch nicht signifikante Veränderungen

⁴ Dass bei einem Leistungszuwachs von rund 33 Punkten an der WBS insgesamt die Leistungssteigerungen in den einzelnen Leistungszügen nicht höher ausgefallen sind, ist u.a. auf die zunehmende Leistungsheterogenität in beiden Leistungszügen zurückzuführen. Die Standardabweichung als Mass für die Streuung der Leistungen um den Mittelwert ist in den beiden Kohorten nach der Umsetzung des «Mathe-Pakets» rund 10 Punkte höher als in den Kohorten 1 und 2 (vgl. Daten im Anhang).

An den meisten WBS-Standorten konnten die Mathematikleistungen gesteigert werden, wobei die Leistungssteigerung oftmals zu gering sind, um statistisch signifikant zu sein. An zwei Standorten sind die Leistungsveränderungen jedoch besonders auffällig. Am Standort A konnten die Mathematikleistungen seit der Einführung des «Mathe-Pakets» um 40 Punkte gesteigert werden (Effektgrösse $d = 0.38$). Am Standort F hingegen sanken die Leistungen um rund 22 Punkte. Gründe, weshalb sich die Leistungen an diesen beiden WBS-Standorten derart unterschiedlich entwickelt haben, können anhand der vorliegenden Daten keine genannt werden. Allgemein sollten die Leistungsveränderungen von kleinen Schülergruppen aber mit Vorsicht interpretiert werden. An den einzelnen WBS-Standorten mit 100 bis 200 Schülerinnen und Schüler pro Testjahr können bereits geringfügige Veränderung der Schülerpopulation grosse Auswirkungen auf die durchschnittlichen Leistungen haben. Zudem wurden seit 2006 einzelne Aufgaben des Mathematiktests verändert und angepasst, was den Schätzfehler bei der Bestimmung der Mittelwerte erhöht. Trotzdem: Eine Leistungssteigerung von 40 Punkten seit der Umsetzung des «Mathe-Pakets» ist ein deutliches Resultat, das nicht allein auf methodische Unsicherheiten zurückgeführt werden kann⁵.

⁵ Bei der PISA-Studie entspricht ein Unterschied von 41 Punkten auf der Mathematikskala – einer Skala mit der gleichen Metrik wie die WBS-Skala – der Leistungsdifferenz zwischen zwei Schuljahren auf der Sekundarstufe I [vgl. OECD (2004). *Lernen für die Welt von morgen. Erste Ergebnisse von PISA 2003*. Paris: OECD. S. 68].

8 Fazit

Im Zentrum der Analysen zum diesjährigen Schlussbericht stehen die Veränderungen der Mathematikleistungen an der WBS. Von besonderem Interesse ist dabei die Frage, ob die Mathematikleistungen durch die unter dem Begriff «Mathe-Paket» subsummierten Massnahmen gesteigert werden konnten.

Die Analysen zeigen, dass die Schülerinnen und Schüler seit der Einführung des «Mathe-Pakets» im Durchschnitt tatsächlich bessere Mathematikleistungen erbringen. Der Leistungsanstieg ist mit rund 14 Punkten jedoch klein. Allerdings wird die beobachtete Leistungssteigerung durch verschiedene demografische und schulpolitische Faktoren entscheidend gedämpft. So wechselte in den Jahren nach der Einführung des «Mathe-Pakets» ein grösserer Anteil der Schülerinnen und Schüler nach der OS ins Gymnasium und ein geringerer Anteil wurde in Kleinklassen und an heilpädagogische Spezialangebote überwiesen. Diese Veränderungen in der Zuweisungspraxis wirken sich negativ auf die Durchschnittsleistungen an der WBS aus. Hätte sich die leistungsrelevante Zusammensetzung der Schülerinnen und Schüler an der WBS nicht verändert, so hätten sich die Mathematikleistungen seit der Einführung des «Mathe-Pakets» um über 30 Punkte verbessert. Ein solcher Leistungsanstieg ist deutlich und kann nicht allein mit zufälligen Leistungsschwankungen zwischen einzelnen Prüfungsjahrgängen erklärt werden.

Die Leistungssteigerung in der Mathematik zeigt sich jedoch nicht bei allen Schülerinnen und Schülern gleichermassen. Während der Anteil Schülerinnen und Schüler mit hohen mathematischen Kompetenzen mit der Umsetzung des «Mathe-Pakets» weiter gesteigert werden konnte, blieb der Anteil Schülerinnen und Schüler mit fehlenden mathematischen Grundkompetenzen konstant. Dies deutet darauf hin, dass in erster Linie gute Schülerinnen und Schüler von den Massnahmen des «Mathe-Pakets» profitieren können. Deutlich wird dies auch, wenn man die Leistungsentwicklung zwischen dem A-Zug und dem E-Zug vergleicht. Im E-Zug konnten die Mathematikleistungen mit dem «Mathe-Paket» statistisch signifikant gesteigert werden. Im A-Zug hingegen haben sich die Leistungen seit der Einführung des «Mathe-Pakets» nicht verändert. Zudem sind die Leistungen im A-Zug sowohl in der Mathematik als auch in Deutsch dieses Jahr über 10 Punkte tiefer als 2010 und damit so tief wie noch nie seit Beginn der Leistungsmessung an der WBS. Diese unterschiedliche Entwicklung in den Leistungen führt zu einer zunehmenden Leistungsheterogenität an der WBS.

Auffällig sind auch die unterschiedlichen Leistungsveränderungen an den sechs WBS-Standorten seit der Einführung des «Mathe-Pakets». An den meisten Standorten konnten keine statistisch signifikanten Leistungssteigerungen festgestellt werden. An einem Standort jedoch haben sich die Mathematikleistungen um 40 Punkte deutlich erhöht. Eine Interpretation und Einordnung dieser Leistungssteigerung muss in erster Linie von der Schulleitung des Standorts sowie von den involvierten Lehrpersonen vorgenommen werden. Interessant wäre es, ob sich aus einer solchen detaillierten Fallanalyse weitere Massnahmen oder Handlungsanleitungen ableiten lassen, die im Sinne von Best Practice von anderen WBS-Standorten übernommen werden könnten.

Anhang

Glossar der statistischen Begriffe

Effektgrösse – Zur Interpretation von statistisch signifikanten Unterschieden wird üblicherweise die Effektgrösse «d» berechnet, indem die Differenz der Mittelwerte durch die Standardabweichungen dividiert wird. Unterschiede, die aufgrund von verschiedenen Skalen zustande gekommen sind, werden so standardisiert und vergleichbar. Eine Effektgrösse von $d = 0.2$ weist auf einen schwachen Unterschied hin, eine Effektgrösse von $d = 0.5$ auf einen mittleren Unterschied und eine Effektgrösse von $d = 0.8$ auf einen starken Unterschied. Auf der standardisierten WBS-Skala können dementsprechend Differenzen von 20 Punkten als klein, solche von 50 Punkten als mittelstark (oder deutlich) und solche von 80 Punkten als sehr gross bezeichnet werden.

Mittelwert – Der Mittelwert entspricht jeweils dem arithmetischen Mittel aller Einzelwerte. Der Mittelwert wird im Bericht mit «M» abgekürzt.

Signifikantes Ergebnis – Ein Ergebnis (Unterschied oder Zusammenhang) ist statistisch signifikant, wenn es durch ein statistisches Testverfahren überprüft und für gültig befunden wurde. Es kann mit einer bekannten, im Voraus festgelegten Irrtumswahrscheinlichkeit α (üblicherweise $\alpha = 0.05$) von der Stichprobe auf die Population geschlossen werden. In diesem Bericht erfüllen signifikante Ergebnisse die Bedingung $p < 0.05$. Die Signifikanz eines Ergebnisses sagt nichts aus über dessen Relevanz. Um signifikante Ergebnisse zu beurteilen, wird daher die Effektgrösse berechnet.

Standardabweichung – Die Standardabweichung ist ein quantitatives Mass für die Streuung der Einzelwerte um den Mittelwert. Entspricht die Verteilung der Einzelwerte einer Normalverteilung, dann besitzt die Standardabweichung die Eigenschaft, dass rund zwei Drittel (68 Prozent) der Einzelwerte zwischen dem Mittelwert \pm eine Standardabweichung $[M \pm 1 \text{ SD}]$ liegen. Wird der Bereich um je eine Standardabweichung erweitert – Mittelwert \pm zwei Standardabweichungen $[M \pm 2 \text{ SD}]$ –, dann befinden sich darin rund 95 Prozent der Einzelwerte. Für die Leistungsdaten auf der standardisierten WBS-Skala ($\text{SD} = 100$ Punkte) heisst das, dass die Ergebnisse von rund zwei Dritteln der Schülerinnen und Schüler zwischen 400 und 600 Punkten und rund 95 Prozent der Ergebnisse zwischen 300 und 700 Punkten liegen. Die Standardabweichung wird im Bericht mit «SD» abgekürzt.

Daten zu den Abbildungen 6.1 bis 6.3

Daten zur Abbildung 6.1: Leistungsentwicklung in den Fächern Mathematik und Deutsch

Mathematik

Testjahr	N	M	SE	SD
2005	759	500	3.39	93
2006	770	511	3.46	96
2007	910	511	3.42	103
2008	897	535	3.57	107
2009	826	506	3.72	107
2010	811	525	3.73	106
2011	777	524	4.12	115

Deutsch

Testjahr	N	M	SE	SD
2005	759	500	3.49	96
2006	770	513	3.64	101
2007	908	525	4.11	124
2008	895	541	3.80	114
2009	825	529	4.00	115
2010	813	527	3.83	109
2011	781	529	4.70	131

Daten zur Abbildung 6.2: Leistungsentwicklung in den Fächern Mathematik und Deutsch im A-Zug

Mathematik

Testjahr	N	M	SE	SD
2005	453	451	3.30	70
2006	360	447	3.85	73
2007	416	440	4.03	82
2008	395	458	4.27	85
2009	387	432	3.81	75
2010	355	451	4.51	85
2011	320	439	5.06	90

Deutsch

Testjahr	N	M	SE	SD
2005	404	445	4.17	84
2006	360	446	4.62	88
2007	414	436	4.29	87
2008	393	450	4.20	83
2009	386	446	4.29	84
2010	355	446	4.58	86
2011	325	430	5.73	103

Daten zur Abbildung 6.3: Leistungsentwicklung in den Fächern Mathematik und Deutsch im E-Zug

Mathematik

Testjahr	N	M	SE	SD
2005	306	572	4.29	75
2006	410	567	3.81	77
2007	494	572	3.47	77
2008	502	596	3.57	80
2009	439	571	4.12	86
2010	456	583	3.88	83
2011	457	584	4.20	90

Deutsch

Testjahr	N	M	SE	SD
2005	355	563	3.52	66
2006	410	572	3.47	70
2007	494	601	4.36	97
2008	502	612	3.50	78
2009	439	601	4.09	86
2010	458	591	3.74	80
2011	456	600	4.65	99